DOCUMENT RESUME

ED 265 520 CS 008 297

AUTHOR Nist, Sherrie L.; Hogrebe, Mark C.
TITLE The Effects of High and Low Relevant Text Underlining on Test Performance.
PUB DATE Dec 85
NOTE 20p.; Paper presented at the Annual Meeting of the National Reading Conference (35th, San Diego, CA, December 3-7, 1985).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Higher Education; *Prior Learning; *Reading Research; *Reading Strategies; Recall (Psychology); *Study Skills; *Test Wiseness; Textbooks
IDENTIFIERS *Outlining

ABSTRACT

A study was conducted to investigate the effects of high and low relevant text underlining and annotating (making margin notes or other marks to indicate important material) on test performance. The subjects, 67 college freshmen, were randomly assigned to one of four experimenter-generated conditions: (1) high relevant underlining; (2) high relevant underlining and annotating; (3) low relevant underlining; or (4) low relevant underlining and annotating, or to a control condition in which students generated their own text markings. During two sessions, all subjects took a test of prior knowledge, read an assigned passage, and took a 24-item multiple choice test consisting of 12 high and 12 low relevant questions. Data analysis showed that subjects in the high relevant condition answered more high relevant questions correctly while low relevant subjects answered more low relevant questions correctly. Subjects who generated their own underlining did not perform significantly better than did those who were given experimenter-generated underlining. The findings show that researcher underlining and annotating have the effect of directing students' attention and are as effective as student markings. However, experimenter markings place different demands on the reader, greatly reducing the number of his or her decisions as to what is or is not important, thus making the reading task easier. (FL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

THE EFFECTS OF HIGH AND LOW RELEVANT TEXT UNDERLINING ON TEST PERFORMANCE

ED265520

Sherrie L. Nist
106 Clark Howell Hall
The University of Georgia
Athens, Georgia 30602

(404) 542-5436

Mark C. Hogrebe
National College of Education

CS008297

Running Head: Text Underlining

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Sherrie L. Nist
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Abstract

The purpose of this study was to investigate the effects of high and low relevant text underlining and annotating on test performance. The subjects were 67 college freshmen who were randomly assigned to one of four experimenter-generated conditions or a control group who generated their own text marking. During two sessions, all subjects took a test of prior knowledge, read the assigned passage, and took a 24-item multiple-choice test consisting of 12 high and 12 low relevant questions. Data analysis showed that subjects in the high relevant group answered more high relevant items correctly while low relevant subjects answered more low relevant items correctly. Subjects who generated their own underlining did not perform significantly better than those who were given experimenter-generated underlining.

The Effects of High and Low Relevant Text Underlining on Test Performance

Text underlining is perhaps the most widely used yet the least researched of all study strategies. One possible explanation for this dilemma is that the process is extremely complex and convoluted, hence difficult to explore. The difficulties of teaching someone how to select the information to be underlined as well as attempting to determine why students chose to underline a specific piece of information present immediate problems for the researcher. Additionally, while college freshmen may have had some instruction in using other study strategies, rarely have they been presented with the opportunity to underline in their texts. It is no wonder, then, that thumbing through a freshmen's psychology text can be a "highlighting" experience--page after page of yellow markings. Freshmen find it inconceivable that their attempts at text interaction will provide little assistance when it comes time to prepare for tests.

One of the reasons why voluminous amounts of underlining are ineffective is due to what is called the von Restorff effect (11). The major tenet of the von Restorff effect is that when an item is isolated against a homogeneous background, increased recall of that item occurs. In the case of underlining, we assume that if text information is underlined, students will tend to study and remember that information better than nonunderlined information. Carrying this logic a step further, if students underlined important information (key ideas, examples, application of theories, etc.), they should remember that information and perform well on tests. If, however, they underlined unimportant information (insignificant details, repetitive text, etc.), they

would tend to study this underlined information and subsequently exhibit poor test performance.

Research investigating text underlining appears to be contradictory in bearing out the above hypotheses. While Crouse and Idstein (5) found that when subjects were given underlined text, they recalled significantly more information than subjects who received nonunderlined text, most research has shown that subjects' overall recall of underlined text does not significantly increase over subjects who received nonunderlined text (6 & 9). However, a finding that is relatively consistent in the extant research is that students who generate their own underlining experience increased recall over those who simply interact with experimenter provided underlining (1, 12, & 9). Greater recall for student generated underlining appears to result from students' elaboration of the material (4) and supports cognitive processing theories of learning (2 & 7). This elaboration assumption is only valid, however, if students are indeed interacting with text while they are underlining in a nonsuperficial manner. If they are using underlining merely as a concentrating technique, hence significantly over underlining, student generated underlining would appear to be of little value. It seems that the true pay-off comes when students use their text markings as a means of reviewing, rehearsing, and preparing for exams. If such is the case, it should make little difference in a controlled research study as to whether the researcher or the subject generates the underlining since the pay-off comes during direct interaction with the underlined material not in the act of the underlining itself. In this sense, the contradictory findings in past underlining/annotating research seem to stem from whether the rearcher was

attempting to add credence to the von Restorff viewpoint or the elaboration processing view to explain the effects of underlining.

The majority of the available research, therefore, seems to be focused on one of two main issues: (1) subject generated versus experimenter generated underlining/annotating and (2) the idea of the von Restorff effect. The present study sought to combine both of these perspectives in an investigation which also examined the effects that prior knowledge might have on an individual's ability to learn from underlined/annotated text. Additionally, most of the other studies used passages that were brief as well as contrived. This study used a rather lengthy text excerpt drawn from a traditional freshman level core course.

Purposes

The present study was designed to answer the following research questions:

1. Is the amount of prior knowledge significantly related to the high relevant, low relevant, or total scores?
2. Do the performances of students who read text material which is either underlined or underlined and annotated differ significantly on multiple-choice questions?
3. Do the performances of students who read text material in which either high or low relevant material is emphasized (i.e. underlined or underlined and annotated) differ significantly on multiple-choice test questions?

4. Is there a significant interaction between the relevancy of emphasis (high or low relevant material emphasized) and the type of emphasis used (underlining or underlining and annotating)?
5. Do the performances of students who read text material that has been underlined and annotated differ significantly from the performances of subjects who read text material with no markings?

Method

Study

The 67 subjects in this study were drawn from the Developmental Studies population at a large southern university. The mean Scholastic Aptitude Test-Verbal score was 370; the mean high school grade point average was 2.7. While not regularly admitted to the university, all subjects possessed intact reading skills (no decoding or severe comprehension problems) but they were generally deficient in processing lengthy text and in their use of study strategies. A series of state mandated and program tests placed all subjects in an upper level reading/study strategies course.

Conditions

All subjects were randomly assigned to one of five conditions

Group 1--High Relevant Underlining (HRU)

Group 2--High Relevant Underlining and Annotating (HRUA)

Group 3--Low Relevant Underlining (LRU)

Group 4--Low Relevant Underlining and Annotating (LRUA)

Group 5--Control, No Underlining or Annotating (C)

For the present study, text aid referred to the two conditions of underlining or underlining and annotating. Annotating refered to making marginal notes or in some way designating important information. Relevancy referred to whether the text aid emphasized high relevant or low relevant material. Groups 1-4 above were the four combinations of text aid and relevancy.

Instruments

The passage was a 2,200 word excerpt from an American government text (3) on the colonization of Anglo-America. The freshman level text was representative of textbooks designed for such college level courses.

To determine high and low relevancy of test material, three experts in reading/studying independently read, underlined, and annotated the passage. Information that was underlined or annotated by two out of three experts was considered high relevant. Such information was considered as important to learn in order to do well on a test covering the chapter. Information that was underlined by only one expert or by no experts was considered low relevant. Low relevant information was that which was considered as unimportant to learn for a test covering the material.

From this expert underlining and annotating, four forms of the passage were devised. Those in the HRU group received the passage with all high relevant information underlined; the HRUA group had high relevant information underlined and annotated. The LRU group received the passage with all low relevant information underlined; the LRUA group had low relevant underlining and low relevant annotation. The Control group received a passage with no underlining or annotating and were instructed to generate their own.

Based on the text excerpt, a 24-item multiple-choice test was constructed which consisted of 12 questions drawn from the high relevant information and 12 questions drawn from the low relevant information.

Finally, a ten-item multiple-choice test of prior knowledge was developed. This test asked general knowledge questions on the colonization of Anglo-America and was used to determine if prior knowledge influenced test performance.

Procedures

The study was carried out in two sessions. During the first session, all subjects were randomly assigned to one of the four experimental conditions or to the control group and were given a folder which contained the ten item test of prior knowledge and the reading passage having one of the five conditions. After taking and handing in the prior-knowledge test, subjects were then given 40 minutes to read and study the passage in preparation for the 24-item multiple-choice test. Those in the control group were encouraged to interact with the passage in any way that would aid in their learning the information. The four experimental conditions were neither encouraged to nor discouraged from additional marking, but were merely told to read and interact with the material in preparation for a multiple-choice test. At the end of the 40 minutes, all folders were collected.

In session two, which occurred the following day, all subjects were given their folder from the previous day and were permitted ten minutes to review. They then took the 24-item multiple-choice test.

Both the prior knowledge test and the test covering the text selection were scored by giving one point for each correct answer. For the text

selection, three raw scores were computed: a high relevant score (number correct out of 12), a low relevant score (number correct out of 12), and the total score (number correct out of 24).

Product-moment correlations were computed to determine the relationship between prior knowledge and high relevant scores, low relevant scores, and total scores. Because of unequal cell sizes, a weighted means analysis of variance was performed for the high relevant, low relevant and total scores (8:116). A Dunnett test (8:300) was used to compare the control group mean with the mean of each condition (HRU, HRUA, LRU, and LRUA).

Results

The means and standard deviations of the five conditions are reported in Table 1 (HRU, HRUA, LRU, LRUA, and Control). The product-moment correlations of prior knowledge with the total score, high relevant questions, and low relevant questions were respectively, $r = .19$, $r = .10$, $r = .19$. These correlation coefficients were not significantly different from zero.

Because of the unequal cell sizes, a weighted means of analysis of variance (8:116) was performed for the high relevant, low relevant and total scores. The results of these analyses are reported in Tables 2, 3, and 4. The main effect for relevancy was significant, $F(1, 49) = 13.56$, $p < .01$ when the dependent measure was the high relevant score. For the low relevant score, $F(1, 40) = 6.98$, $p < .05$. The main effect for text aid and all interactions between text aid and relevancy were not significant for the three dependent measures.

A Dunnett test (8:33) was used to compare the control group mean with the mean of each experimental condition. This test was employed in order to keep

the alpha level of .05 constant for the family of comparisons defined by each dependent variable. For each dependent variable, the mean of the control group did not differ significantly from the most disparate mean of the four conditions.

Discussion

This study sought to examine both the role of elaboration and the von Restorff effect as they relate to text marking. The results of the present study suggested that when text material was underlined or underlined and annotated, students directed their attention to the information emphasized by these text aids and supports the operation of the von Restorff effect. Students who read the passage in which high relevant material was emphasized by a text aid, correctly answered more high relevant questions than students who read the passage in which low relevant material was emphasized. Likewise, students who read the passage in which low relevant material was emphasized correctly answered more low relevant questions than students who read the passage with emphasis on high relevant material. These results suggested that text material which had been underlined and annotated by the researchers had a rather strong influence on directing the reader's attention to certain parts of the text. The emphasized text material appeared to be processed more thoroughly, in that performance on questions covering the emphasized text were significantly better than on questions covering the unemphasized sections of the passage. This finding was particularly interesting because it indicated that the reader's attention was influenced more by the author's underlining and annotating than by what were, in reality, the relevant and important parts of the passage. Students who read text that had low relevant material

underlined had more difficulty in determining that the underlined material was unimportant and, subsequently, were unable to redirect their attention to the more important ideas and concepts in the passage.

This effect of relevancy of the emphasized text material was a rather strong one. Eta-squared, which is a measure of the amount of variance in the dependent variable accounted for by the treatment variable, was .21 for the score on the high relevant questions and .12 for the score on the low relevant questions. Rosenthal and Rubin (10) offered a way of interpreting the effect of a treatment called the "binomial effect size display" (BESD). An eta-squared of .21 indicated that 73% of the students in the group in which the high relevant material was emphasized performed above the median score on the high relevant questions, while only 27% of the students in the low relevant group performed above the median score on the high relevant questions. Similarly, when the score on the low relevant questions was used as the dependent variable, an eta-squared of .12 indicated that 67% of the subjects in the low relevant group performed above the median compared to 33% of the students in the high relevant group.

The results of the analyses in which the performances of the control group were compared to the performances of the text aid-relevancy groups suggested that when underlining and annotation were not supplied by the author, students were able to answer as many questions correctly as students in the treatment groups but not significantly more. This finding seems to suggest that the students in the control group attempted to remember as much of the passage as they could, with little discrimination about whether the information had high or low relevance. The significant results for relevancy discussed previously suggested that students in the treatment groups directed

their attention toward the underlined and annotated material. Since there were no significant differences on the dependent variables between the control group and any of the treatment groups, the results indicate that subjects who were allowed to use their own reading and marking techniques did not make clear distinctions between high and low relevant information and did not elaborate information with anymore depth than those who received experimenter generated underlining. This finding differs from the bulk of previous research which found that students who generated their own underlining outperformed those who received researcher generated underlining (9). Evidently, students in the control group did a fairly effective job of learning both high and low relevant material because the performances of students in conditions which emphasized either high or low relevant material were no better than the performances of the control subjects.

Annotation of the passage did not increase test performance over underlining alone. Experimenter supplied underlining was sufficient in directing students' attention to the high or low relevant material. This finding, however, does not offer any insight about how student generated annotation might improve comprehension and recall.

Prior knowledge was not significantly related to the number of questions answered correctly. The mean prior knowledge score was 4.27 out of 10 questions with a standard deviation of 1.6. This rather low mean score on prior knowledge of the passage content, together with the small standard deviation suggested that students in the sample had relatively little prior knowledge on the topic. We assume that the nonsignificance and little variability on the prior knowledge variable stems partially from the fact that there was also little variation in verbal abilities of these subjects as

measured by SATV scores. Hence, we would expect that the amount of prior knowledge would also show little variation.

The results of this investigation indicated that researcher underlining and annotating have an effect of directing students' attention, and are equally effective as student generated underlining. But experimenter supplied markings place substantially different demands on the reader than do reader generated markings. Experimenter generated markings greatly reduce the number of decisions for the reader as to what is important and, in a sense, make the student's task easier. The present study offers no insight about the processes that students in the control group employed in deciding what they should remember. If the control group subjects adopted a "remember all" strategy, then they were allocating their mental resources inefficiently. Subsequent research needs to focus on the strategies students use in making decisions about what is important and what should be marked. Nevertheless, the results of the present study do suggest that students tend to key in on and learn what is underlined, whether the information is relevant or irrelevant, and lends credence to the von Restorff effect as an explanation as to why underlining can be a successful study strategy. While the underlining in this study was experimenter generated, it is safe to generalize that when students do their own underlining that they also tend to learn that which they underline.

References

1. Bobrow, S. A. & Bower, G. H. Comprehension and recall of sentences. Journal of Experimental Psychology, 1963, 80:455-461.

2. Bransford, J. D. Human cognition. Belmont, CA: Wadsworth Publishing Company, 1979.

3. Burns, J. M., Peltason, J. W., & Cronin, T. E. Government by the people. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1978.

4. Craik, F. I. M., & Tulving, E. Depth of processing and the retention of words in episodic memory. Journal of Experimental Psychology: General, 1975, 104:268-294.

5. Crouse, J. H., & Idstein, P. Effects of encoding cues on prose learning. Journal of Educational Psychology, 1972, 63:309-313.

6. Fowler, F. L., & Barker, A. S. Effectiveness of highlighting for retention of text material. Journal of Applied Psychology, 1974, 59:358-364.

7. Jenkins, J. J. Four parts to remember: A tetrahedral model of memory experiments. In L. S. Cermak & F. I. M. Craik (Eds.), Levels of processing in human memory. Hillsdale, NJ: Erlbaum, 1979.

8. Myers, J. L. Fundamentals of experimental design (3rd ed.). Boston: Allyn & Bacon, Inc., 1979.

9. Rickards, J. P., & August, G. J. Generative underlining strategies in prose recall. Journal of Educational Psychology, 1975, 67:860-865.

10. Rosenthal, R., & Rubin, D. B. A simple, general purpose display of magnitude of experimental effect. Journal of Educational Psychology, 1982, 74:166-169.

11. Wallace, W. P. Review of the historical, empirical, and theoretical status of the von Restorff phenomenon. Psychological Bulletin, 1965, 63:410-424.

12. Wittrock, M. C. Learning as a gererative process. Educational Psychology, 1974, 11, 87-95.

TABLE I

MEANS AND STANDARD DEVIATIONS FOR MULTIPLE-CHOICE TEST SCORES

	TOTAL SCORE		HIGH RELEVANT SCORE		LOW RELEVANT SCORE	
	M	SD	M	SD	M	SD
Low Relevant Material Underlining Only (n = 13)	13.54	5.78	5.92	2.40	7.62	3.55
Low Relevant Material Underlining & Annotation (n = 14)	15.29	2.64	6.71	1.90	8.57	1.45
High Relevant Material Underlining Only (n = 14)	14.57	3.01	8.07	1.59	6.50	1.79
High Relevant Material Underlining & Annotation (n = 12)	15.00	3.49	8.75	2.22	6.25	.05
Control No Text Aid (n = 13)	16.23	2.01	8.08	1.61	8.15	1.46

TABLE 2

WEIGHTED MEANS ANALYSIS OF VARIANCE FOR THE TOTAL SCORE

Source	df	SS	MS	F
Text aid	1	1.1836	1.1836	1.00
Relevancy	1	.1396	.1396	.12
Interaction	1	.4346	.4346	.36
Error	49		1.1829	

TABLE 3

WEIGHTED MEANS ANALYSIS OF VARIANCE FOR THE HIGH RELEVANT SCORE

Source	df	SS	MS	F
Text Aid	1	.5401	.5401	1.6732
Relevancy	1	4.3765	4.3765	13.5579**
Interaction	1	.0031	.0031	.01
Error	49		.3228	

** $p < .01$

TABLE 4

WEIGHTED MEANS ANALYSIS OF VARIANCE FOR THE LOW RELEVANT SCORE

Source	df	SS	MS	F
Text Aid	1	.1246	.1246	.29
Relevancy	1	2.9529	2.9529	6.9759*
Interaction	1	.3636	.3636	.86
Error	49		.4233	

* $p < .05$